

## A Derivations

### A.1 Equation (3)

$$\begin{aligned}
\mathbf{M}_i \mathbf{M}'_i &= Cov(\mathbf{X}_i) - \sigma_i^2 \mathbf{I}_T = \mathbf{U}_i \mathbf{D}_i \mathbf{U}'_i - \sigma_i^2 \mathbf{U}_i \mathbf{U}'_i \\
&= \mathbf{U}_{i1} \mathbf{D}_{i1} \mathbf{U}'_{i1} + \mathbf{U}_{i2} \mathbf{D}_{i2} \mathbf{U}'_{i2} - \sigma_i^2 (\mathbf{U}_{i1} \mathbf{U}'_{i1} + \mathbf{U}_{i2} \mathbf{U}'_{i2}) \\
&= \mathbf{U}_{i1} \mathbf{D}_{i1} \mathbf{U}'_{i1} + \mathbf{U}_{i2} \sigma_i^2 \mathbf{I}_{T-Q_i} \mathbf{U}'_{i2} - \mathbf{U}_{i1} (\sigma_i^2 \mathbf{I}_{Q_i}) \mathbf{U}'_{i1} - \mathbf{U}_{i2} (\sigma_i^2 \mathbf{I}_{T-Q_i}) \mathbf{U}'_{i2} \\
&= \mathbf{U}_{i1} (\mathbf{D}_{i1} - \sigma_i^2 \mathbf{I}_{Q_i}) \mathbf{U}'_{i1} \\
&= \left[ \mathbf{U}_{i1} (\mathbf{D}_{i1} - \sigma_i^2 \mathbf{I}_{Q_i})^{1/2} \right] \left[ \mathbf{U}_{i1} (\mathbf{D}_{i1} - \sigma_i^2 \mathbf{I}_{Q_i})^{1/2} \right]' \\
&= \left[ \mathbf{U}_{i1} (\mathbf{D}_{i1} - \sigma_i^2 \mathbf{I}_{Q_i})^{1/2} \mathbf{A}_i \right] \left[ \mathbf{U}_{i1} (\mathbf{D}_{i1} - \sigma_i^2 \mathbf{I}_{Q_i})^{1/2} \mathbf{A}_i \right]',
\end{aligned}$$

where  $\mathbf{A}_i$  is an arbitrary rotation matrix.

### A.2 Exact EM algorithm

#### A.2.1 Full expected log likelihood

$$\begin{aligned}
Q_1(\Theta | \hat{\Theta}^{(k)}) &= \sum_{v=1}^V E \left[ \log g(\mathbf{y}_i(v); \mathbf{A}_i \mathbf{s}_i(v), \nu_0^2 \mathbf{C}_i) | \mathbf{y}_i(v), \hat{\Theta}^{(k)} \right] \\
&= -\frac{QV}{2} \log(\nu_0^2) - \frac{V}{2} \sum_{q=1}^{Q_i} \log(c_q) - \frac{1}{2\nu_0^2} \sum_{v=1}^V \left\{ \mathbf{y}_i(v)' \mathbf{C}_i^{-1} \mathbf{y}_i(v) - 2\mathbf{y}_i(v)' \mathbf{C}_i^{-1} \mathbf{A}_i E[\mathbf{s}_i(v) | \mathbf{y}_i(v), \hat{\Theta}^{(k)}] \right. \\
&\quad \left. + \text{Tr} \left( \mathbf{A}_i' \mathbf{C}_i^{-1} \mathbf{A}_i E[\mathbf{s}_i(v) \mathbf{s}_i(v)' | \mathbf{y}_i(v), \hat{\Theta}^{(k)}] \right) \right\},
\end{aligned}$$

$$\begin{aligned}
Q_2(\Theta | \hat{\Theta}^{(k)}) &= \sum_{v=1}^V E \left[ \log g(\mathbf{s}_{i1}(v); \mathbf{s}_0(v), \Sigma_v) | \mathbf{y}_i(v), \hat{\Theta}^{(k)} \right] \\
&= -\frac{1}{2} \sum_{v=1}^V \sum_{q=1}^L \log \nu_q^2(v) - \frac{1}{2} \sum_{v=1}^V \left\{ \text{Tr} \left( \Sigma_v^{-1} E[\mathbf{s}_{i1}(v) \mathbf{s}_{i1}(v)' | \mathbf{y}_i(v), \hat{\Theta}^{(k)}] \right) \right. \\
&\quad \left. - 2\mathbf{s}_0(v)' \Sigma_v^{-1} E[\mathbf{s}_{i1}(v) | \mathbf{y}_i(v), \hat{\Theta}^{(k)}] + \mathbf{s}_0(v)' \Sigma_v^{-1} \mathbf{s}_0(v) \right\} \\
&= -\frac{1}{2} \sum_{v=1}^V \sum_{q=1}^L \left\{ \log \nu_q^2(v) + \frac{1}{\nu_q^2(v)} \left( E[s_{iq}^2(v) | \mathbf{y}_i(v), \hat{\Theta}^{(k)}] - 2s_{0q}(v) E[s_{iq}(v) | \mathbf{y}_i(v), \hat{\Theta}^{(k)}] + s_{0q}^2(v) \right) \right\}
\end{aligned}$$

$$\begin{aligned}
Q_3(\Theta|\hat{\Theta}^{(k)}) &= -\frac{1}{2} \sum_{v=1}^V \sum_{q=L+1}^{Q_i} E \left[ \log g(s_{iq}(v); \mu_{iq, z_{iq}(v)}, \sigma_{iq, z_{iq}(v)}^2) | \mathbf{y}_i(v), \hat{\Theta}^{(k)} \right] \\
&= -\frac{1}{2} \sum_{v=1}^V \sum_{q=L+1}^{Q_i} \sum_{m=1}^M Pr(z_{iq}(v) = m | \mathbf{y}_i(v), \hat{\Theta}^{(k)}) E \left[ \log g(s_{iq}(v); \mu_{iq, z_{iq}(v)}, \sigma_{iq, z_{iq}(v)}^2) | z_{iq}(v) = m, \mathbf{y}_i(v), \hat{\Theta}^{(k)} \right] \\
&= -\frac{1}{2} \sum_{v=1}^V \sum_{q=L+1}^{Q_i} \sum_{m=1}^M Pr(z_{iq}(v) = m | \mathbf{y}_i(v), \hat{\Theta}^{(k)}) \left\{ \log \sigma_{iqm}^2 + \frac{1}{\sigma_{iqm}^2} \left( E[s_{iq}(v)^2 | z_{iq}(v) = m, \mathbf{y}_i(v), \hat{\Theta}^{(k)}] \right. \right. \\
&\quad \left. \left. - 2\mu_{iqm} E[s_{iq}(v) | z_{iq}(v) = m, \mathbf{y}_i(v), \hat{\Theta}^{(k)}] + \mu_{iqm}^2 \right) \right\}
\end{aligned}$$

$$Q_4(\Theta|\hat{\Theta}^{(k)}) = \sum_{v=1}^V \sum_{q=L+1}^{Q_i} E \left[ \log \pi_{iq, z_{iq}(v)} | \mathbf{y}_i(v), \hat{\Theta}^{(k)} \right] = \sum_{v=1}^V \sum_{q=L+1}^{Q_i} \sum_{m=1}^M Pr(z_{iq}(v) = m | \mathbf{y}_i(v), \hat{\Theta}^{(k)}) \log \pi_{iqm}.$$

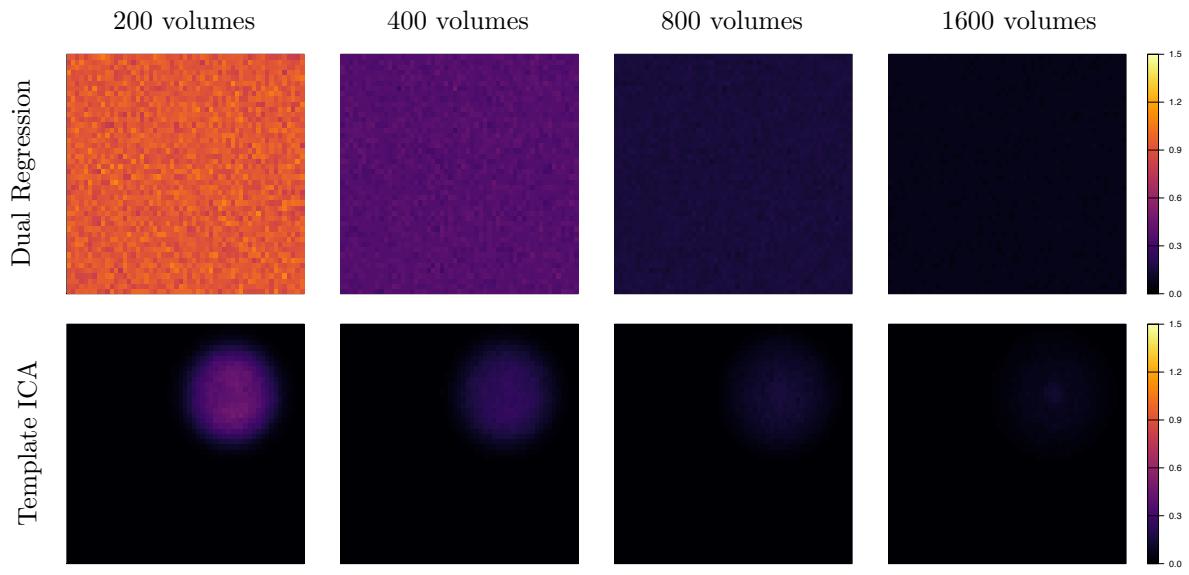
### A.2.2 Conditional posterior distribution of $\mathbf{s}_i(v)$

$$\begin{aligned}
p(\mathbf{s}_i(v) | \mathbf{z}_i(v), \mathbf{y}_i(v), \Theta) &\propto p(\mathbf{y}_i(v) | \mathbf{s}_i(v), \mathbf{z}_i(v), \Theta) p(\mathbf{s}_i(v) | \mathbf{z}_i(v), \Theta) \\
&= p(\mathbf{y}_i(v) | \mathbf{s}_i(v), \mathbf{z}_i(v), \Theta) p(\mathbf{s}_{i1}(v) | \mathbf{z}_i(v), \Theta) p(\mathbf{s}_{i2}(v) | \mathbf{z}_i(v), \Theta) \\
&= g(\mathbf{y}_i(v); \mathbf{A}_i \mathbf{s}_i(v), \nu_0^2 \mathbf{C}_i) g(\mathbf{s}_{i1}(v); \mathbf{s}_0(v), \boldsymbol{\Sigma}_v) g(\mathbf{s}_{i2}(v); \boldsymbol{\mu}_{\mathbf{z}_i(v)}, \mathbf{D}_{\mathbf{z}_i(v)}) \\
&\propto g(\mathbf{s}_i(v); \boldsymbol{\mu}_{\mathbf{s}|\mathbf{z}, \mathbf{y}}(v), \boldsymbol{\Sigma}_{\mathbf{s}|\mathbf{z}, \mathbf{y}}(v))
\end{aligned}$$

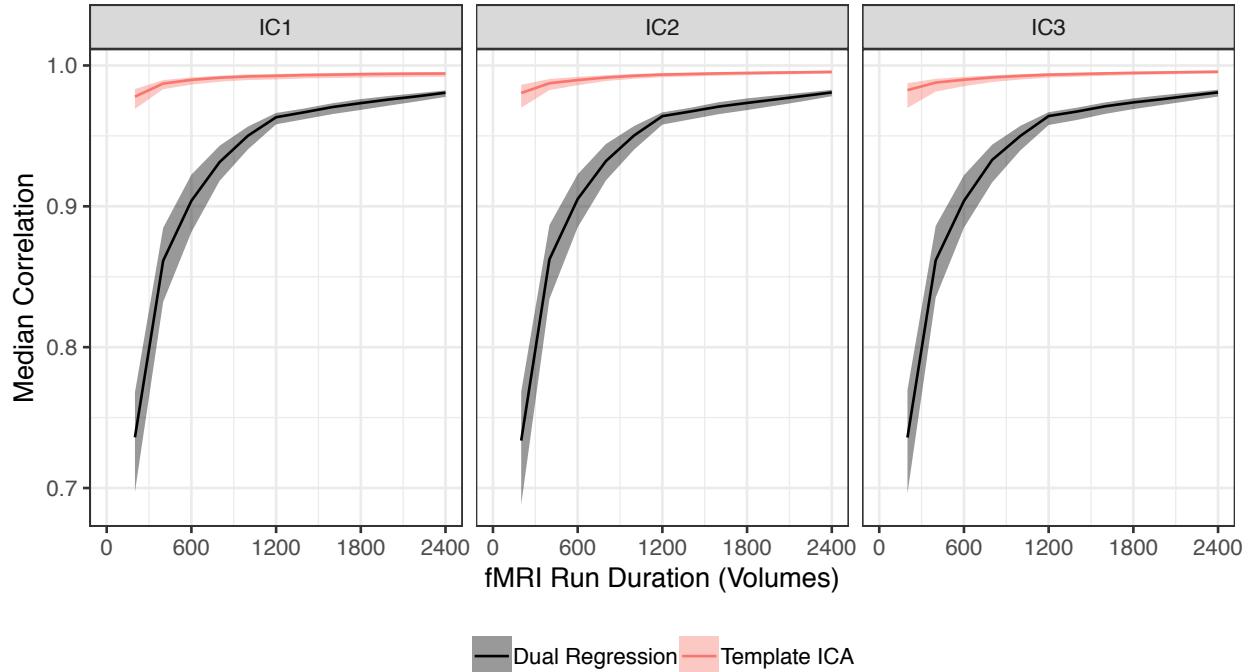
### A.2.3 Posterior distribution of $\mathbf{z}_i(v)$

$$\begin{aligned}
p(\mathbf{z}_i(v) | \mathbf{y}_i(v), \Theta) &= \frac{p(\mathbf{y}_i(v) | \mathbf{z}_i(v), \Theta) p(\mathbf{z}_i(v) | \Theta)}{p(\mathbf{y}_i(v) | \Theta)} \\
&= \frac{p(\mathbf{y}_i(v) | \mathbf{z}_i(v), \Theta) p(\mathbf{z}_i(v) | \Theta)}{\sum_{\mathbf{z}_i(v) \in \mathcal{R}_i} p(\mathbf{y}_i(v) | \mathbf{z}_i(v), \Theta) p(\mathbf{z}_i(v) | \Theta)} \\
&= \frac{g(\mathbf{y}_i(v) : \boldsymbol{\mu}_{\mathbf{y}|\mathbf{z}}(v), \boldsymbol{\Sigma}_{\mathbf{y}|\mathbf{z}}(v)) \prod_{q=L+1}^{Q_i} \pi_{iq, z_{iq}(v)}}{\sum_{\mathbf{z}_i(v) \in \mathcal{R}_i} \left\{ g(\mathbf{y}_i(v) : \boldsymbol{\mu}_{\mathbf{y}|\mathbf{z}}(v), \boldsymbol{\Sigma}_{\mathbf{y}|\mathbf{z}}(v)) \prod_{q=L+1}^{Q_i} \pi_{iq, z_{iq}(v)} \right\}},
\end{aligned}$$

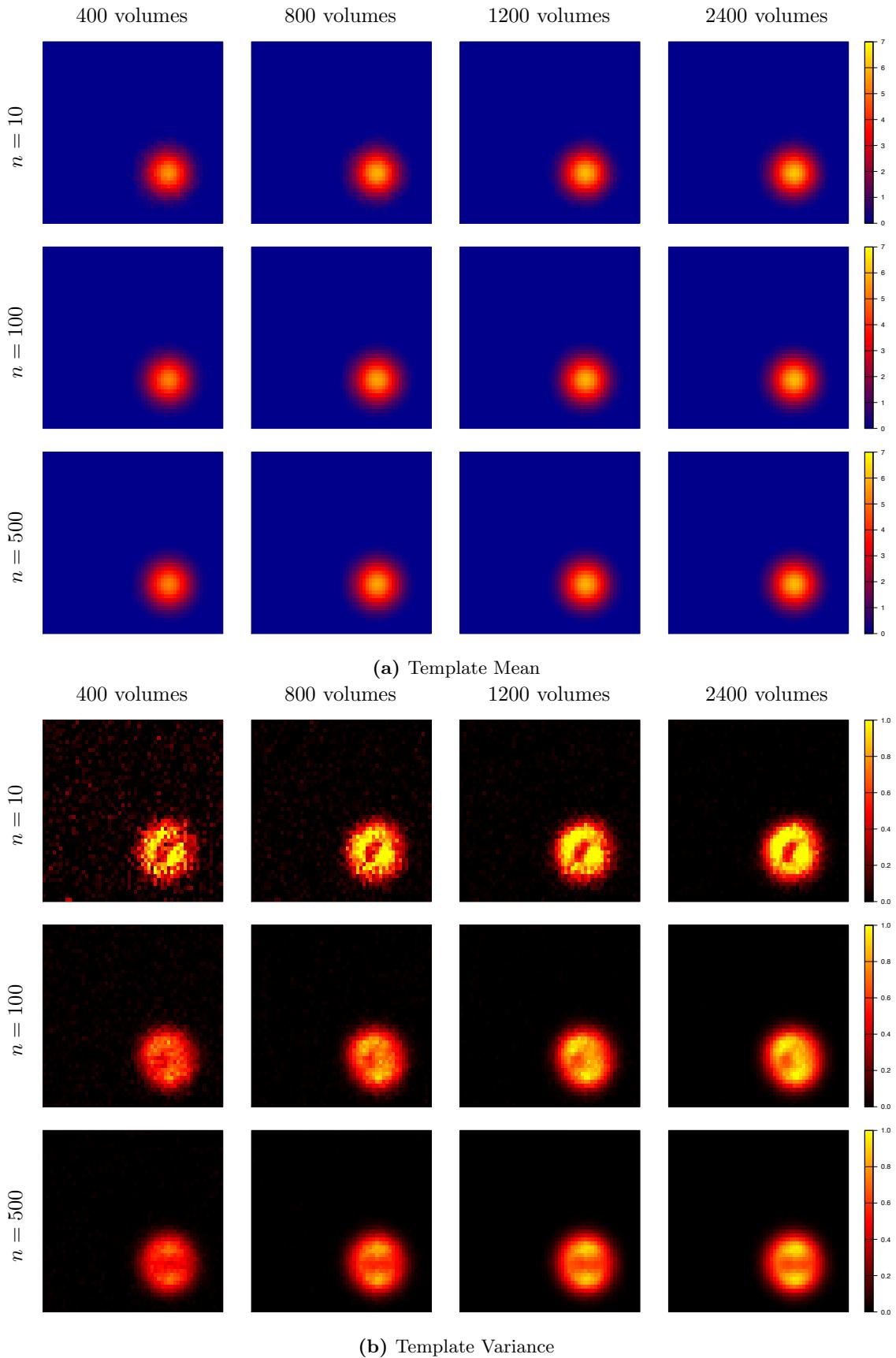
## B Simulation B Figures



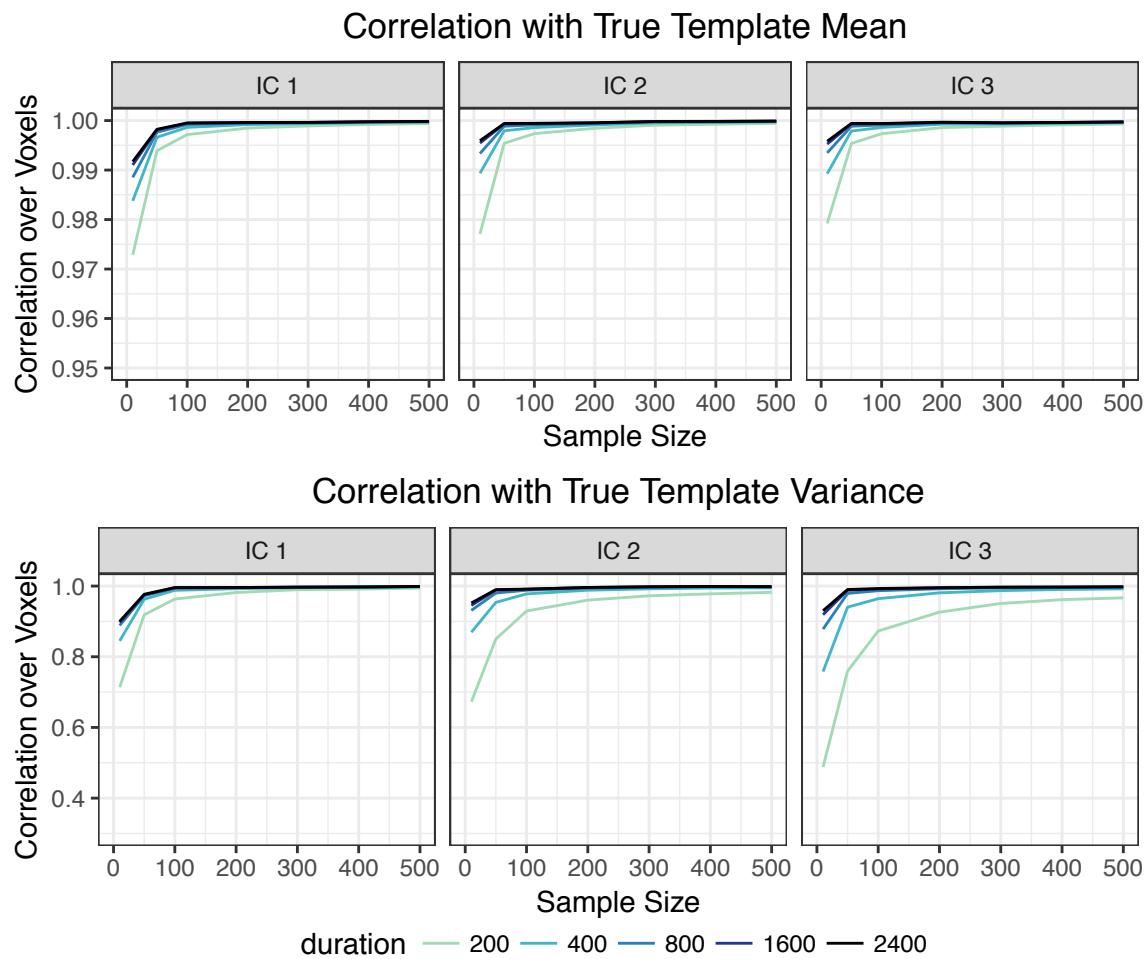
**Figure A1:** MSE across subjects of dual regression and template ICA estimates versus the ground truth for one source signal in Simulation B.



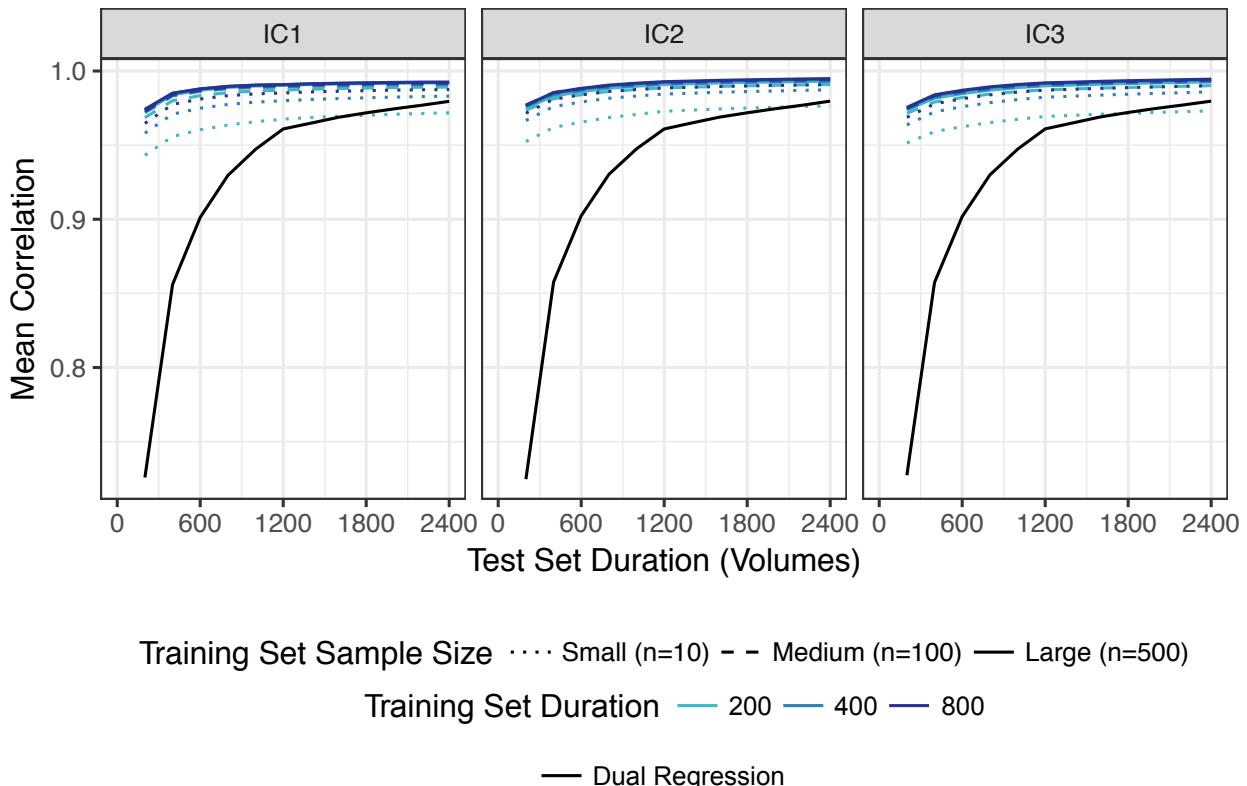
**Figure A2:** Correlation between the true and estimated source signals across all voxels activated at the group level in Simulation B. Lines represent the median across all subjects, and shaded ribbons represent the first and third quartiles.



**Figure A3:** Template estimates by sample size and scan duration for one source signal in Simulation B.



**Figure A4:** Correlation between the true and estimated templates in Simulation B.



**Figure A5:** Correlation between the true and estimated source signals across all voxels activated at the group level in Simulation B, averaged over subjects, by scan duration of subjects in the test set. Color indicates the scan duration of subjects in the training set (used for template estimation), and line type indicates the sample size of the training set.